

Boundary Conditions for Limited-Area Ensemble Kalman Filters

Ryan D. Torn¹ & Gregory J. Hakim

University of Washington, Seattle, WA

Chris Snyder²

National Center for Atmospheric Research, Boulder, CO

Submitted to *Monthly Weather Review*

January 5, 2005

¹*Corresponding Author:*

Ryan D. Torn
Department of Atmospheric Sciences, Box 351640
University of Washington
Seattle, WA 98195-1640
E-mail: torn@atmos.washington.edu

²The National Center for Atmospheric Research is sponsored by the National Science Foundation

Abstract

One aspect of implementing a limited-area ensemble Kalman filter (EnKF) involves the specification of a suitable ensemble of lateral boundary conditions. We propose two classes of methods to populate a boundary condition ensemble. In the first class, the ensemble of boundary conditions is provided by an EnKF on a larger domain and is approximately a random draw from the probability distribution function for the forecast (or analysis) on the limited-area domain boundary given the available observations. The second class perturbs around a deterministic estimate of the state using assumed spatial and temporal covariance relationships. Methods in the second class are relatively flexible and easy to implement.

Experiments that test the utility of these methods are performed for both an idealized low-dimensional model and limited-area simulations using the Weather Research and Forecasting (WRF) model; all experiments employ simulated observations under the perfect model assumption. The performance of the ensemble boundary condition methods is assessed by comparing the results of each experiment against a control “global” EnKF that extends beyond the limited-area domain. For all methods tested, results show that errors for the limited-area EnKF are larger near the lateral boundaries than those from a control EnKF, but decay into the limited-area domain so that errors are comparable to the control case. The relatively larger errors in the limited-area domain EnKF originate from not assimilating observations outside the limited-area domain and, in the second class of methods, from deficiencies in boundary spatial and temporal covariances. Overall, these experiments suggest that for observation densities typical in NWP, ensemble boundary conditions can be specified in the absence of a global ensemble without significant penalty in the domain interior by perturbing around an ensemble mean.

1. Introduction

The ensemble Kalman filter (EnKF) provides a computationally practical alternative to the extended Kalman filter for use in atmospheric data assimilation (e.g. Evensen 1994, Evensen 2003, Hamill 2005). One attractive property of the EnKF is state-dependent estimates of the background-error covariance matrix, which controls how observational information affects model variables. This attribute may be especially important for mesoscale data assimilation, where the synoptic-scale balance relationships (geostrophic and hydrostatic balance), employed in many existing data assimilation schemes (Parrish and Derber 1992, Rabier et al. 1998), are weak or absent. While previous studies do indeed suggest that the EnKF will be useful for mesoscale applications (Snyder and Zhang 2003, Dowell et al. 2004), these studies did not account for lateral boundary condition error; accounting for such error may be crucial for situations with inflow across domain boundaries. Here we address this issue of ensemble boundary conditions for limited-area EnKFs by exploring a range of practical options.

Previous work on limited-area modeling has shown that lateral boundary conditions affect solutions over the entire domain (Warner et al. 1997). Predictability on limited-area domains is distinct from global domains in that growing errors may be advected out of the domain, which along with accurate boundary conditions on inflow boundaries may lead to enhanced predictability relative to global domains (Errico and Baumhefner 1987). Conversely, sensitivity studies have shown that errors on limited-area domains may be traced back to the boundaries (Errico et al. 1993), and therefore inaccurate boundary conditions may result in large forecast errors (Vukicevic and Errico 1990, Vukicevic and Paegle 1989, Warner et al. 1989). In general, the size of the domain and the length of model integration determine how much the lateral boundaries affect solutions on the domain interior. Moreover, Nutter et al. (2004a, 2004b) show that ensemble boundary conditions must be perturbed to prevent an ensemble forecast from losing variance for increasing lead time.

Boundary condition error has not been addressed in past EnKF studies. Most previous EnKF studies have employed global models that resolve synoptic scales (e.g. Houtekamer

and Mitchell 2001, Mitchell et al. 2002, Whitaker and Hamill 2002, Whitaker et al. 2004). EnKF studies with limited-area models have used identical boundary conditions for each ensemble member (Snyder and Zhang 2003, Dowell et al. 2004, Zhang et al. 2005). Here, we explore a range of possible methods for ensemble boundary conditions in a limited-area EnKF for two different models in ideal settings. Although this paper deals with ensemble data assimilation, we note that the boundary condition ideas outlined here may also prove useful for limited-area ensemble forecasting.

This paper is organized as follows. Section 2 outlines five methods for generating an ensemble of lateral boundary conditions. Section 3 details the implementation of these methods in EnKFs applied to a simple one-dimensional model and the Weather Research and Forecast model (WRF). We document the results of the experiments for the simple model in section 4 and for the WRF experiments in section 5. A concluding summary is provided in section 6.

2. Boundary Perturbation Methods

A common practice in numerical weather prediction models is to advance the lateral boundary points linearly in time, such that $X_{k,i}$, a state variable at time t_k and location i on the boundary, evolves according to:

$$X_{k+1,i} = m\Delta t + X_{k,i}, \tag{1}$$

where

$$m = \frac{X_{k+1,i} - X_{k,i}}{\Delta t_{bc}},$$

and Δt_{bc} is the time between boundary updates (t_{k+1} and t_k). Whereas $X_{k,i}$ comes from the limited-area analysis, values of $X_{k+1,i}$ are typically based on forecasts from a global model or a model with a larger domain.

As summarized previously, the specification of lateral boundary conditions is a key component of limited-area models. Since lateral boundary conditions are imperfect and affect the solution on the domain interior, an ensemble forecast in a limited-area model

requires an ensemble of lateral boundary conditions. This ensemble of boundary values, $X_{k+1,i}^j$, where j is the ensemble member index, must take into account spatial and temporal covariance relationships to avoid spurious effects in the interior, while maintaining ensemble variance.

We define two classes of ensemble boundary conditions, the first of which derives completely from an ensemble on a larger domain. The absence, in many applications, of an appropriate larger domain ensemble motivates the second class, which consists of methods that parameterize or model the spatial and temporal covariance relationships of the boundary conditions. Below we define one implementation of the first class, and four methods in the second class. This range of possible methods is not meant to be exhaustive, but rather to provide a broad sample of techniques that emphasize the available flexibility in the problem. Table 1 provides a summary of the essential aspects of the methods that we now discuss in detail.

a. Global Ensemble Values (GEV)

Perhaps the most natural method for ensemble lateral boundary conditions is one that uses values from a global-model ensemble having the correct short-term error statistics, such as might come from an EnKF. This method involves pairing each limited-area ensemble member with a global ensemble member. Appealing to (1), $X_{k+1,i}^j$ is obtained by interpolating the parent global ensemble member's short-term forecast onto the boundary points of the limited-area domain. The main advantage of this method is that the covariance relationships are state dependent. A disadvantage is that errors in, and even the formulation of, the global model can lead to detrimental incompatibilities on the boundaries; for example, the global and limited-area model may have different attractors or the global ensemble could be drawn from the wrong distribution. Furthermore, the global ensemble must have at least as many members as desired for the limited-area ensemble.

Presently such an operational global EnKF system does not exist. Even assuming a global EnKF system becomes available, retrospective case studies may require running a global ensemble if an archived one does not exist. It is therefore advantageous to seek

other methods that yield results comparable to GEV, but that are not as computationally expensive and which offer greater flexibility.

Next, we introduce a series of methods that populate an ensemble around a deterministic forecast. In each of the following methods, $X_{k+1,i}^j$ is calculated from

$$X_{k+1,i}^j = \bar{X}_{k+1,i} + \delta X_{k+1,i}^j, \quad (2)$$

where $\bar{X}_{k+1,i}$ is the ensemble mean boundary value, which is given by the best available estimate of boundary points (e.g. from an operational center analysis or forecast) and $\delta X_{k+1,i}^j$ is a perturbation for ensemble member j generated by one of the following methods.

b. Climatology Time Series (CTS)

CTS obtains values for $\delta X_{k+1,i}^j$ from a set of time series taken from a climatological dataset. Time series equal in number to the ensemble size are drawn from climatology starting at arbitrary dates for a length of time equal to the period of assimilation; each time series includes all the boundary points and is assigned to a specific ensemble member. The ensemble mean is removed from the ensemble and these perturbations, $\delta X_{k+1,i}^j$, are scaled by an assumed constant. Because this method does not require a separate, independent global ensemble, it is less computationally intensive than GEV.

Spatial and temporal covariances in this case are not state dependent but rather they estimate climatological covariances to within sampling errors, and therefore should reflect the dominant balance relationships such as geostrophic or hydrostatic balance. To the extent that balances are significantly non-linear, scaling the perturbations will disrupt the balance. These attributes are both a strength and a weakness because CTS yields time series with spatial and temporal covariances from real atmospheric states, but they are flow independent. Another drawback is the empirical scaling of the boundary perturbations.

c. Global Ensemble Sampling (GES)

It is possible to choose a few boundary points to perturb and to adjust all other boundary points to maintain an assumed balance. We define two methods (GES and, in section 2d,

LAES) that have assumed spatial and temporal covariance relationships, but do not require an empirical scaling factor for the boundary variance. GES consists of choosing boundary points to model as an autoregressive process, adjusting all other boundary points to maintain balance with each of the new perturbed boundary points using an assumed fixed covariance matrix, then summing up the perturbations over all chosen points. Mathematically, this process can be described via

$$\delta X_{k+1,i}^j = \sum_{m=1}^{N_{pert}} \left[(a\delta X_{k,m}^j + \sqrt{1-a^2}\sigma_{k,m}\epsilon_{k,m}^j C_{i-m}) \right], \quad (3)$$

where a is the temporal autocorrelation coefficient for Δt_{bc} and N_{pert} is the total number perturbed boundary points. $\delta X_{k,m}^j$ is ensemble member j 's deviation from the ensemble mean, C_{i-m} is the assumed covariance between $X_{k+1,i}$ and $X_{k+1,m}$, $\sigma_{k,m}$ is the ensemble standard deviation and $\epsilon_{k,m}^j$ are independent realizations of a normally distributed random variable having zero mean and unit variance.

In the present method, $\sigma_{k,m}$ is taken to be the sample standard deviation from a global ensemble at the m th grid point. If a global ensemble exists with fewer members than needed, this method may be used to augment the ensemble for use in the GEV method; here we consider the case where the global ensemble is simply sampled for its variance to populate a random ensemble. Similarly, the global ensemble could be used to estimate C_{i-m} at each t_k , but we will only consider the case that C_{i-m} is fixed and independent of t_k . This method is attractive because it uses a state-dependent variance estimate, but may suffer from sampling errors in that estimate and from deficiencies in the assumed temporal autocorrelation coefficient and spatial covariances C_{i-m} , which are state independent.

d. Limited-Area Ensemble Sampling (LAES)

This method is similar to GES, but uses the standard deviation given by the limited-area ensemble analysis on the boundary points as the source of $\sigma_{k,m}$ in (3). One defect of this method is that the ensemble variance decreases too much with time owing to sampling error in EnKF; therefore, the limited-area ensemble boundary variance is inflated by a small

amount. It is also not clear that the limited-area ensemble analyses reflect the uncertainty in the global fields because they lack information about observations and error evolution outside the domain. An advantage of this method is that it is completely self-contained, requiring no external data. Like the previous method, the use of assumed spatial and temporal covariances is a weakness.

e. Fixed Covariance Perturbations (FCP)

Unlike the previous two methods, which only directly perturb a limited number of boundary points, this method employs fixed spatial and temporal covariance relationships to create perturbations in all boundary points and fields. Here, we assume that the boundary perturbations are random draws from a normal distribution with zero mean and covariance matrix P^b , which we shall denote by $N(O, P^b)$. The temporal evolution of the boundary perturbations is assumed to be an autoregressive process such that,

$$\delta X_{k+1,i}^j = a\delta X_{k,i}^j + \sqrt{1 - a^2}\tilde{X}_{k,i}^j, \quad (4)$$

where a is the temporal autocorrelation coefficient for Δt_{bc} and $\tilde{X}_{k,i}^j$ is a scaled sample from $N(O, P^b)$. In general, it is nontrivial to construct and sample from P^b for systems with a high-dimensional multivariate state. Fortunately, variational assimilation systems typically include sophisticated models for P^b and are often available in a form which makes sampling from $N(O, P^b)$ straightforward (e.g. Andersson et al. 2000, Barker et al. 2004). One of the main advantages of this method is that balanced perturbations can be created without storing large amounts of data as may be required for GEV and CTS; however, this method may suffer from a lack of state-dependent information.

3. Experiments

a. Simple Model

A simple one-dimensional model is used to explore the boundary condition methods described above in a controlled setting (Lorenz and Emanuel 1998, hereafter LE). This model has proven useful in EnKF experiments because it qualitatively relates to mid-latitude

dynamics and has low computational cost (e.g. Anderson 2001; Whitaker and Hamill 2002). At each grid point, i , the model is defined by

$$\frac{dX_i}{dt} = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F, \quad (5)$$

where the terms on the righthand-side can be thought of as representing advection, dissipation and forcing, respectively, of an unspecified, non-dimensional meteorological quantity. Long integrations of this model yield a time-mean value for X_i of 2.3 and variance of 13.

We define a “global” version of the model using 80 grid points and periodic boundary conditions (that is $X_{81} = X_1$), and a “limited-area” version of the model using 50 grid points and specified lateral boundary conditions. Global and limited-area grid points are co-located, eliminating the need for interpolation. Inspection of (5) reveals that the limited-area domain contains three grid points that must be specified, two on the left boundary and one on the right ($i=1, 2, 50$); these boundary points are interpolated linearly in time via (1).

All experiments are performed in a perfect-model context with truth taken as a long integration from a randomly generated initial state for a total duration of 10,000 assimilation times. Sixty global ensemble members are created by randomly drawing from model climatology and limited-area members are initialized from the global ensemble. Each assimilation time is 0.05 model time units long; one model time unit can be thought of as 120 hours (Lorenz and Emanuel, 1998). Observations are taken from the truth state with added Gaussian random error having zero mean and a variance of 0.5; observation errors are assumed uncorrelated. The global observation network consists of 40 observations at each time t_k , which are located in between every other grid point [$y_{k,i} = \frac{1}{2}(X_{k,2i} + X_{k,2i+1}), i = 1, 2, 3, \dots, 40$]. For the limited-area EnKF, only the 25 observations located within the limited-area domain are assimilated.

One of the drawbacks of EnKFs is spurious long-distance covariances due to under-sampling, which yield erroneous corrections to the background field (Hamill et al. 2001, Houtekamer and Mitchell 1998, Houtekamer and Mitchell 2001). To overcome this problem, observation influence is localized during the assimilation using the Gaspari and Cohn (1999)

fifth-order piecewise rational function given by their equation (4.10); the covariances reach zero 25 grid points from the observation location. We employ the ensemble square-root filter so that observations are unperturbed (Whitaker and Hamill 2002).

Spurious covariances also artificially reduce the variance of the ensemble, thus producing an over-confident estimate of the state (Hamill et al. 2001). Anderson (1999) showed that inflating the ensemble perturbations by a small amount can overcome this aspect of undersampling. We choose the covariance inflation factor so that the ratio of the RMS error of the global ensemble mean (E1) to the average RMS error of each global ensemble member (E2) is 0.71, as expected for a 60 member ensemble (Murphy 1988). The resulting inflation factor (1.007) is employed for all LE experiments. Details regarding the implementation of the boundary condition methods are summarized in Table 2.

b. WRF Model

The experiments performed with the Weather Research and Forecasting (WRF) model are designed to mimic the simple model experiments as closely as possible. Since WRF is a limited-area model, we create a pseudo-global domain (PGD) with boundaries well removed from the limited-area domain (LAD)¹. Boundary points are linearly interpolated in time by (1).

The chosen domain has 100 km grid spacing over 90 grid points in each horizontal direction and 27 vertical levels (Fig. 1). This relatively coarse horizontal grid is employed to test the boundary methods on scales with well-known dynamics and covariance relationships (e.g. approximately geostrophic and hydrostatic). The LAD (50 grid points in each horizontal direction) shares the PGD grid and vertical levels such that interpolation is not required (Fig. 1).²

¹Well-removed is defined here such that solutions on the LAD boundaries are unaffected by the location of the PGD boundaries.

²All model runs employ the MRF boundary layer scheme, Betts-Miller-Janjic convective scheme and no cloud microphysics. On the PGD, a lateral boundary zone three grid points wide provides a transition from specified time behavior to free time behavior; the outer boundary point is specified by interpolated GFS values and the inner two boundary points by a linear combination of interpolated GFS values and WRF

All experiments are performed under the perfect-model assumption for 15 days with observation assimilation every six hours. Initial and boundary conditions for the truth run are derived from National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS) analyses every six hours starting at 18 UTC 18 March 2003³.

Whitaker et al. (2004) showed that surface pressure observations are sufficient to constrain the tropospheric mass field using an EnKF in a sparse observation network. Thus, for simplicity, we assimilate only surface pressure (technically, column integrated dry air mass) every six hours for an idealized observation network. The global observation network is a randomly located, uniformly distributed collection of 250 grid points, 72 of which are located within the LAD (Fig. 1). Additive observation errors are uncorrelated and Gaussian with zero mean and a variance of 1.0 hPa.

As for the LE experiments, ad hoc steps are taken to overcome the effect of ensemble undersampling. Covariance localization is again accomplished by the Gaspari and Cohn compact correlation function that reduces the covariances to zero at a radius of 3000 km from an observation. Global and limited-area ensemble covariance inflation factors of 1.06 and 1.075, respectively, are determined by calibrating the system by the ratio of the RMS error in the ensemble mean to the average RMS error in each ensemble member for the dry-air mass field.

A 100-member PGD ensemble is initialized with 100 GFS lagged forecasts validating within 12 hours of 18 UTC 18 March 2003. An assimilation step is performed prior to the first forecast in order to reduce the ensemble variance; the LAD ensemble is initialized by a similar procedure. To generate ensemble boundary conditions for the PGD, the CTS method is employed using time series of GFS analyses and an empirically derived scaling factor of 0.125. Details regarding the implementation of the boundary condition methods for the limited-area domain are summarized in Table 2.

dynamics.

³This period was chosen because of the rich variety of weather systems (cutoff cyclones, ridges, strong surface cyclones) that affected the area of interest

4. Simple Model Results

Three diagnostics are used to evaluate the different boundary condition methods. The RMS error over space and time is defined as

$$E = \sqrt{\frac{1}{N_t} \frac{1}{N_x} \sum_{k=1}^{N_t} \sum_{i=1}^{N_x} [\bar{X}_{k,i} - X_{k,i}^t]^2}. \quad (6)$$

The RMS error as a function of time is defined as

$$E_k = \sqrt{\frac{1}{N_x} \sum_{i=1}^{N_x} [\bar{X}_{k,i} - X_{k,i}^t]^2}, \quad (7)$$

and the RMS error as a function of space is defined as

$$E_i = \sqrt{\frac{1}{N_t} \sum_{k=1}^{N_t} [\bar{X}_{k,i} - X_{k,i}^t]^2}. \quad (8)$$

$\bar{X}_{k,i}$ is the ensemble mean at time t_k and location i , $X_{k,i}^t$ is the corresponding value of truth and N_x and N_t are, respectively, the number of grid points and number of times.

Fig. 2 shows that after a spin-up of 50 assimilation times, the global ensemble RMS error in the simple model is stable in time, exhibiting only short-lived deviations from the time-mean value of E_k (0.30), which is 60 % less than the observation error. This error distribution indicates that assimilating 40 observations leads to smaller errors in the global ensemble state than just interpolating the observations alone. For clarity of presentation, RMS errors for the other methods are not shown in Fig. 2. As expected, the global ensemble has the smallest RMS error over space and time, but all boundary methods have RMS errors within 5 % of the global ensemble's value as seen in Table 3.

Whereas the global ensemble has uniform errors in space due to the periodicity of the model and uniformity of the observation network, the limited-area domain results show larger errors near the lateral boundaries. More precisely, Fig. 3 shows the differences in the RMS error as a function of space between each limited-area domain boundary-method experiment and the global ensemble ($E_i^{method} - E_i^{global}$); positive values indicate where errors

are larger than the global ensemble on the limited-area domain. Error differences in the limited-area domain interior are indistinguishable from the global ensemble, but positive differences exist near the boundaries. A few of the methods actually have error differences that are slightly negative, but this may be due to the limited number of analyses used to compute the statistics. The increased error near the boundaries results from two sources other than linear time interpolation: fewer observations assimilated near lateral boundaries, and boundary covariance assumptions. Daley (1992) showed that assimilating observations with a Kalman filter in a periodic domain with a data void leads to larger errors near the interface between the areas with sparse and dense observations, similar to these experiments where the limited-area domain does not assimilate observations outside the domain. GEV has the smallest errors among all limited-area domain boundary methods since it has state-dependent boundary covariances. The differences in RMS error between the global ensemble and limited-area ensembles are reduced by assimilating with a denser observation network. When re-running each experiment with twice as many observations, the maximum error differences near the boundaries are reduced by 30 % for the methods that use assumed covariances as compared to Fig. 3 (not shown).

Recall from section 2 that we defined two classes of methods, one that uses state dependent statistics and another that uses assumed statistics. Even though CTS and FCP differ in part by their treatment of the temporal evolution of the boundary perturbations, these two methods have similar error (Fig. 3). This suggests that the temporal behavior of the boundary conditions can be approximated well by an autoregressive model. In supplementary experiments, we varied the autocorrelation coefficient and scaling parameter for the FCP method. These experiments suggest that for small scaling parameters (less than 0.5; i.e. small ensemble variance), the error differences are similar for all autocorrelation values used; however, for larger scaling parameters (≥ 0.80 ; i.e. larger ensemble variance), using the default autocorrelation coefficient ($a = 0.94$) lead to 20 % smaller error differences near the left boundary as compared to $a = 0$ (not shown).

GES has the smallest errors among the methods that use assumed covariance relation-

ships. This is due to the fact that the state-dependent ensemble variance used in this method scales the ensemble perturbations consistent with the state, which yields smaller forecast errors near the boundaries as compared to the CTS and FCP methods, which have fixed, state-independent variance.

Even though GES and LAES use similar procedures, they yield different results because of the variance estimate ($\sigma_{k,m}$) used for (3). We assume that the ratio of the 0.05 time-unit forecast variance to the analysis variance is constant (1.05) when inflating the limited-area ensemble variance; however, this ratio actually ranges from 0.5-3.5. As such, the LAES ensemble boundary conditions will often have the wrong estimate of ensemble variance, leading to larger forecast errors during subsequent integrations. Furthermore, the autocorrelation coefficient is important for both GES and LAES; error differences double when white noise ($a = 0$) is assumed.

In summary, these results based on LE indicate that: as expected, the state-dependent covariances for boundary conditions from a “global” EnKF (GEV) produce the smallest errors, although all methods have RMS errors over space and time that are within 5 % of the global ensemble method. Moreover, an autoregressive model appears to provide a good approximation for the temporal evolution of the boundary perturbations, and a variance estimate from a global ensemble is preferable to scaling ensemble boundary perturbations with a constant scaling factor.

5. WRF Results

We proceed to evaluate the performance of the boundary condition methods in a WRF EnKF, with the following considerations. Unlike the one-dimensional LE model which has only three boundary points, WRF has four two-dimensional lateral boundaries; covariance relationships among boundary points are therefore an important concern. For these experiments, the model is dry and the model state is represented by six prognostic variables (u, v, w , temperature, height, and dry-air mass). Two different metrics are evaluated: error in the dry-air mass (observed), and error in the 500 hPa geopotential height (unobserved);

other tropospheric fields give similar results, including vertical motion. We note that the observation network used here is sparse when compared to the LE experiments. The ratio of the number of degrees of freedom to number of observations for LE is 2:1; whereas for the WRF experiments the ratio is approximately 5000:1.

Fig. 4 shows that the PGD RMS error as a function of time for both dry-air mass and 500 hPa height exhibits no obvious trend; therefore, the 250 surface pressure observations constrain the state. The RMS error in dry-air mass (Fig. 4a) is approximately 0.7 hPa after 3 days, which is 30 % smaller than the actual observation errors, thus indicating that assimilating observations on 3 % of the dry-air mass grid points leads to smaller error than just simply interpolating the observations. Table 4 shows that the PGD ensemble has the lowest RMS error over space and time in the dry-air mass field although all methods have RMS errors within 35 % of the PGD value. Although mid-tropospheric observations are not assimilated, 500 hPa height errors are steady in time, which is consistent with the results of Whitaker et al. (2004). Similar to the dry-air mass field errors, all boundary methods have 500 hPa geopotential height RMS errors over space and time within 25 % of the PGD value (Table 4). Moreover, all boundary condition methods have qualitatively similar plots to both panels of Fig. 4 and are not shown. To infer the impact of observations, we integrated the ensemble for 15 days without assimilating observations. The results of this experiment reveal that the assimilation of surface pressure observations leads to approximately a factor of three decrease in the RMS error in both fields (not shown).

The impact of observations on analysis error is illustrated by the spatial structure of the RMS error for the PGD run (Fig. 5). Generally, errors are smaller close to observation sites and greater in observation network gaps, notably over the central United States, Saskatchewan and eastern Mexico. The areas of largest error are also co-located with grid points that have the highest variance about the 15-day time-mean. Significant errors over eastern Mexico are related to the dry-air mass covariances, which tend to be strong for places of common elevation, but small otherwise, so that observations on the Mexican Plateau and the Gulf of Mexico yield small corrections along the sloping model topography.

In contrast to the dry-air mass field, errors in the 500 hPa geopotential height field (Fig 5b) reflect the mean path of significant weather systems during this period, rather than the observation network. Areas of largest errors are co-located with grid points that experience persistent shortwave-trough passages and have the highest variance in 500 hPa geopotential height about the 15-day time-mean.

RMS error differences between the PGD and LAD boundary-condition experiments (Fig. 6a) support the previous LE results by indicating that the RMS error in the dry-air mass is larger near the lateral boundaries, particularly near large gaps in the observation network (e.g. Hudson Bay and Vancouver Island). Along each of the four lateral boundaries, errors are approximately 0.5 hPa larger than for the PGD; just a few points into the interior however, errors are similar to those for the PGD. All methods have qualitatively similar RMS error differences compared to the PGD, so only FCP is shown in this figure. Note that these results are not exclusive to surface pressure observations. A similar error-difference profile is obtained when assimilating temperature observations at three different vertical levels at the same sites as the dry-air mass observations (not shown).

We create a composite that takes into account the RMS error as a function of distance from the boundary, similar to Fig. 3a for LE. The composite in Fig. 6b is produced by averaging the RMS error differences over all grid points as a function of distance from the lateral boundary. In general, errors are larger near the lateral boundaries, but are nearly indistinguishable from the PGD field 500 km into the domain interior (~ 5 grid points). Fig 6b also suggests that assumed boundary covariances and state-dependent covariances have similar error for dry-air mass. This contrasts with LE results which showed that methods with state-dependent covariances have smaller error near the boundaries. Methods with assumed covariances may have RMS errors similar to methods with state-dependent covariances for WRF because the dominant covariance length scale is relatively large compared to the grid spacing. Whereas LE's grid points are corrected by a few observations because the covariance length scale is limited to a few grid points, WRF grid points are corrected by more observations because the covariance length scales are larger

and observations are distributed in two dimensions.

Similar to Fig. 6a, Fig. 7a shows the RMS error differences between FCP and PGD for the 500 hPa geopotential height field. RMS errors are larger for FCP along the boundaries, although the largest errors are located in the interior of the domain over Michigan, which may be due to a small number of samples for these experiments. Specifically, the errors over Michigan can be traced to errors in two amplifying shortwave troughs in strong northwesterly flow. This region of increased error only exists for methods that use assumed boundary covariances, suggesting that the northern boundary's assumed covariances did not accurately model the error as the trough entered the domain. The second area of larger errors near Hudson Bay is associated with an eastward propagating trough along the northern boundary; although this trough is resolved on the PGD, it is treated by a combination of linear tendencies and model dynamics on the LAD due to its proximity to the lateral boundary.

Constructing an analogous composite to Fig. 6b for 500 hPa geopotential height (Fig. 7b) also shows that the error differences decline toward the domain interior; however, more significant differences exist among the methods near the boundaries in this field as compared to the dry-air mass, indicating that some covariance assumptions lead to smaller errors than others for unobserved fields. The following discussion will focus on comparing the various methods in class two.

Figs. 6b and 7b show that RMS errors are smaller for the FCP method as compared to the CTS method. The differences in temporal and spatial covariances contribute to this result. One might expect the FCP spatial covariances to be better than the CTS spatial covariances since the FCP covariance model is based on forecast differences, while the CTS covariance model is based on differences from states drawn from climatology. Similar to the LE model experiments, the error in FCP is found to be relatively insensitive to the choice of autocorrelation coefficient (not shown).

Contrary to LE results, Figs. 6b and 7b suggest that even though GES uses a state-dependent variance to scale boundary perturbations, FCP and CTS have smaller error when applied to a WRF EnKF. As noted in Table 2, we directly perturb the model sigma-level 0.737

temperature on the boundary; therefore, only this field has boundary perturbations with explicit state-dependent variance. Boundary perturbations in all other fields are calculated to maintain balance with the prescribed temperature perturbations, thus the variance in these other fields may not match the global ensemble’s estimate. The results presented here are sensitive to the field chosen to perturb; errors are larger in an experiment where we perturb only the boundary dry-air mass instead of temperature. LAES has larger error than GES for both fields since the temperature variance in LAES is scaled by the limited-area analysis variance, which may not reflect the external forecast uncertainty like the global ensemble variance used in GES.

6. Discussion and Conclusions

We have presented several methods for creating an ensemble of lateral boundary conditions for use in a limited-area EnKF. The most obvious ensemble boundary conditions come from an EnKF run outside the limited-area domain, and this method yields the lowest errors. Since such an ensemble may often be unavailable, or of the wrong size or resolution, a range of alternative boundary conditions have been explored. The results are encouraging because they suggest that errors for the alternative methods are only modestly larger than for the control case, and that these errors are limited to an error “boundary layer” near the edges of the limited-area domain. Moreover, the alternative boundary conditions are relatively computationally inexpensive compared to the cost of the EnKF on a larger domain, and they offer a range of flexibility in specifying the boundary conditions. The performance of these methods has been demonstrated in a simple one-dimensional model (LE) and in the WRF model.

The magnitude and horizontal distribution of the boundary errors depend on the covariance assumptions related to each method, with state-dependent covariances leading to reduced boundary errors, especially for unobserved fields. For example, the control case (GEV) had the lowest errors among all methods for both LE and WRF, which demonstrates the positive impact of state-dependent covariances near the boundaries. Unfortunately, most

operational centers are not currently producing short-range ensemble forecasts with the appropriate statistical properties for the GEV method; therefore current implementations of this method require computing an expensive global ensemble for use in the limited-area EnKF.

Fortunately, our experiments also showed that there exists a second class of computationally inexpensive methods based on assumed spatial and temporal covariances for lateral boundary conditions that may provide a useful alternative to the GEV method. One such method, CTS, employs ensemble perturbations generated by scaling deviations from randomly drawn climatological time series. Similar to all methods in this second class, errors for CTS were comparable to GEV for observed fields, but were larger for unobserved fields near the boundaries. Again, this suggests the importance of state-dependent cross-variable covariances. Moreover, since the boundary perturbation scaling is arbitrary, the boundary variance is not state dependent, which may result in forecast errors with the wrong variance.

Among the methods in the second class with assumed spatial covariances, the fixed boundary covariance method, FCP, had the smallest errors. Boundary perturbations in this case were cheaply generated by drawing from a specified multivariate Gaussian distribution, such as the covariance model for the WRF 3D-VAR algorithm. In this method, the state-independent spatial covariance used on the boundaries comes from forecast differences and are representative of short-term forecast errors (Rabier et al. 1998), thus they are probably more appropriate than random climatological perturbations. As with CTS, a potential drawback of this method is that boundary perturbations are scaled empirically. It may be possible to choose the scaling factor adaptively at each assimilation time based on the statistics of the innovation vector.

A generalization of the fixed-covariance method, GES, employs a large-scale ensemble only to estimate the ensemble variance. This method is attractive because of its low computational cost, and ability to use a global ensemble with fewer members than desired on the limited-area domain. Only a subset of the degrees of freedom are directly perturbed on the boundaries, and an assumed covariance model is used to perturb all other points.

We note that although the perturbed points have the variance of the global ensemble, the adjusted points in general will not. Experiments presented here show that the quality of the analyses varied depending on the choice of points; thus deciding what fields to perturb is problematic. Finally, a variant on the GES method, LAES, uses scaled estimates of ensemble variance from the limited-area ensemble *analyses*. This method tends to have larger errors than GES because the variance of the forecast boundary perturbations originate from the limited-area analyses, which may not be representative of forecast errors near the lateral boundaries.

Two of the above boundary methods can be combined using a coarse grid nesting approach resulting in state-dependent boundary covariances at a reasonable computational cost. On the coarse (outer) grid, a limited-area ensemble may be forced by a method from the second class outlined here. Assimilation of observations on this domain produces state-dependent forecast covariances, and allows the use of the best method (GEV) on the finer (inner) nests. In theory, this procedure should result in state-dependent covariances on the fine-scale domain boundaries, without the cost of running a true global ensemble. Furthermore, this approach also allows for observations outside the inner nested domain to affect the state in the inner domain because an ensemble of observation estimates is available from the coarse grid. Another potential option for situations with sparse observations near the lateral boundaries involves linearly combining the limited-area analysis ensemble mean with a deterministic global analysis, similar to the relaxation zone procedure used for forecasts on the boundaries of limited-area models. Future work will focus on testing these ideas in mesoscale EnKF applications.

Acknowledgments.

We thank Dale Barker and Mi-Seon Lee for help with the WRF 3D-VAR system. This work was supported by NSF Grant ITR-0205648 and NOAA CSTAR Grant NA17RJ1232.

References

- Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.
- Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- Andersson, E., M. Fisher, R. Munro and A. McNally, 2000: Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation scheme, and the explanation of a case of poor convergence. *Quart. J. Roy. Meteor. Soc.*, **126**, 1455–1472.
- Barker, D. M., W. Huang, Y.-R. Guo, and Q. N. Xiao, 2004: A three-dimensional (3DVAR) data assimilation system for use with MM5: implementation and initial results. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Daley, R., 1992: Forecast-error statistics for homogeneous and inhomogeneous observation networks. *Mon. Wea. Rev.*, **120**, 627–643
- Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005.
- Errico, R. M., and D. P. Baumhefner, 1987: Predictability experiments using a high resolution limited-area model. *Mon. Wea. Rev.*, **114**, 1625–1641.
- Errico, R. M., T. Vukicevic, and K. Raeder, 1993: Comparison of initial and boundary condition sensitivity for a limited-area model. *Tellus*, **45A**, 539–557.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model

- using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162.
- Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance- dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.
- Hamill, T. M., 2005: Ensemble-based atmospheric data assimilation: A tutorial. In press.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: simulations with a small model. *J. Atmos. Sci.*, **55**, 399–414.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 89–125.
- Nutter, P., D. Stensrud, and M. Xue, 2004a: Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble

- forecasts *Mon. Wea. Rev.*, **132**, 2358–2377.
- Nutter, P., M. Xue, and D. Stensrud, 2004b: Application of lateral boundary condition perturbations to help restore dispersion in limited-area ensemble forecasts. *Mon. Wea. Rev.*, **132**, 2378–2390.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center’s spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Rabier, F., J.-N. Thepaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **124**, 1–39.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.
- Vukicevic, T., and J. Paegle, 1989: The influence of one-way interacting boundary conditions upon predictability of flow in bounded numerical models. *Mon. Wea. Rev.*, **117**, 340–350.
- Vukicevic, T., and R. M. Errico, 1990: The influence of artificial factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev.*, **118**, 1460–1482.
- Warner, T. T., L. E. Key, and A. M. Lario, 1989: Sensitivity of mesoscale model forecast skill to some initial data characteristics: Data density, data position analysis procedure and measurement error. *Mon. Wea. Rev.*, **117**, 1281–1310.
- Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **78**, 2599–2617.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed

observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**, 1190–1200.

Zhang, F., Z. Meng, and A. Aksoy, 2005: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. *Mon. Wea. Rev.*, in review.

Figure Captions

Figure 1. WRF pseudo-global domain (full grid), limited-area domain (inner box) and surface pressure observation network (dots). Shading indicates model surface elevation every 600 m.

Figure 2. RMS error as a function of time (E_k) for the LE global ensemble.

Figure 3. Difference between each method's RMS error as a function of space and global ensemble's RMS error as a function of space ($E_i^{meth} - E_i^{global}$) for all limited-area domain grid points. Positive values indicate RMS errors greater than the global ensemble.

Figure 4. (a) RMS error as a function of time (E_k) for the WRF PGD dry-air mass field over all grid points in the limited-area domain. (b) As in (a) for 500 hPa geopotential height.

Figure 5. (a) RMS error as a function of space (E_i) in the WRF PGD dry-air mass field. Colorbar units are hPa. (b) As in (a) for 500 hPa geopotential height. Colorbar units are meters. Dots indicate observation locations.

Figure 6. (a) Difference between the RMS error as a function of space in dry-air mass for the FCP method and the RMS error as a function of space for the PGD dry-air mass field at corresponding grid points in hPa. Shading indicates positive differences (i.e. larger errors than PGD), while contours denote negative differences in 0.1 hPa increments. Dots are observation locations. (b) Composite RMS error differences in each boundary method's dry-air mass field as a function of distance from the boundary.

Figure 7. As in Fig. 6, but for 500 hPa geopotential height. Units are meters.

Table 1: General characteristics of each ensemble boundary condition method.

	Global Ensemble Values	Climatology Time Series	Global Ensemble Sampling	Limited-Area Ensemble Sampling	Fixed Covariance Perturbations
Acronym	GEV	CTS	GES	LAES	FCP
spatial covariance	state-dependent ensemble	climatology	fixed, assumed	fixed, assumed	fixed, assumed
temporal covariance	state-dependent ensemble	individual time series	autoregressive model	autoregressive model	autoregressive model
perturbation scaling	implicit	scaling parameter	global ensemble variance	limited-area ensemble variance	scaling parameter

Table 2: Details regarding the implementation of ensemble boundary condition methods of the second class in LE and WRF.

	LE Model	WRF Model
Climatology Time Series (CTS)		
time series source	random integrated states	sample from 1 year of GFS analyses
scaling	0.09 (time series variance = global ensemble)	0.11 (time series variance = global ensemble)
Global Ensemble Sampling (GES)		
perturbation points	every boundary point	temperature at level 0.737 at every fourth point
variance source	global ensemble	global ensemble
spatial covariance	time-averaged covariance	WRF 3D-VAR NCEP covariance
autocorrelation	0.94 (from global ensemble)	0.55 (from global ensemble)
Limited-Area Ensemble Sampling (LAES)		
perturbation points	every boundary point	temperature at level 0.737 at every fourth point
variance inflation	1.05 (from global ensemble)	1.00 (inflation worsened results)
spatial covariance	time-averaged covariance	WRF 3D-VAR NCEP covariance
autocorrelation	0.94 (from global ensemble)	0.55 (from global ensemble)
Fixed Covariance Perturbations (FCP)		
spatial covariance source	time averaged covariance	WRF 3D-VAR NCEP covariance
scaling	0.40 (increment variance = global ensemble)	0.65 (increment variance = global ensemble)
autocorrelation	0.94 (from global ensemble)	0.55 (from global ensemble)

Table 3: RMS error over space and time (E) for each boundary method with the LE model.

Boundary Method	RMS Error (E)
Global Ensemble	0.30
Global Ensemble Values	0.31
Climatology Time Series	0.32
Fixed Covariance Perturbations	0.31
Global Ensemble Sampling	0.31
Limited-Area Ensemble Sampling	0.32

Table 4: RMS error over space and time (E) for each boundary method and each metric with the WRF model.

Boundary Method	dry-air mass (hPa)	500 hPa height (m)
Global Ensemble	0.63	9.58
Global Ensemble Values	0.68	10.16
Climatology Time Series	0.72	11.69
Fixed Covariance Perturbations	0.72	10.00
Global Ensemble Sampling	0.78	11.31
Limited-Area Ensemble Sampling	0.83	11.45

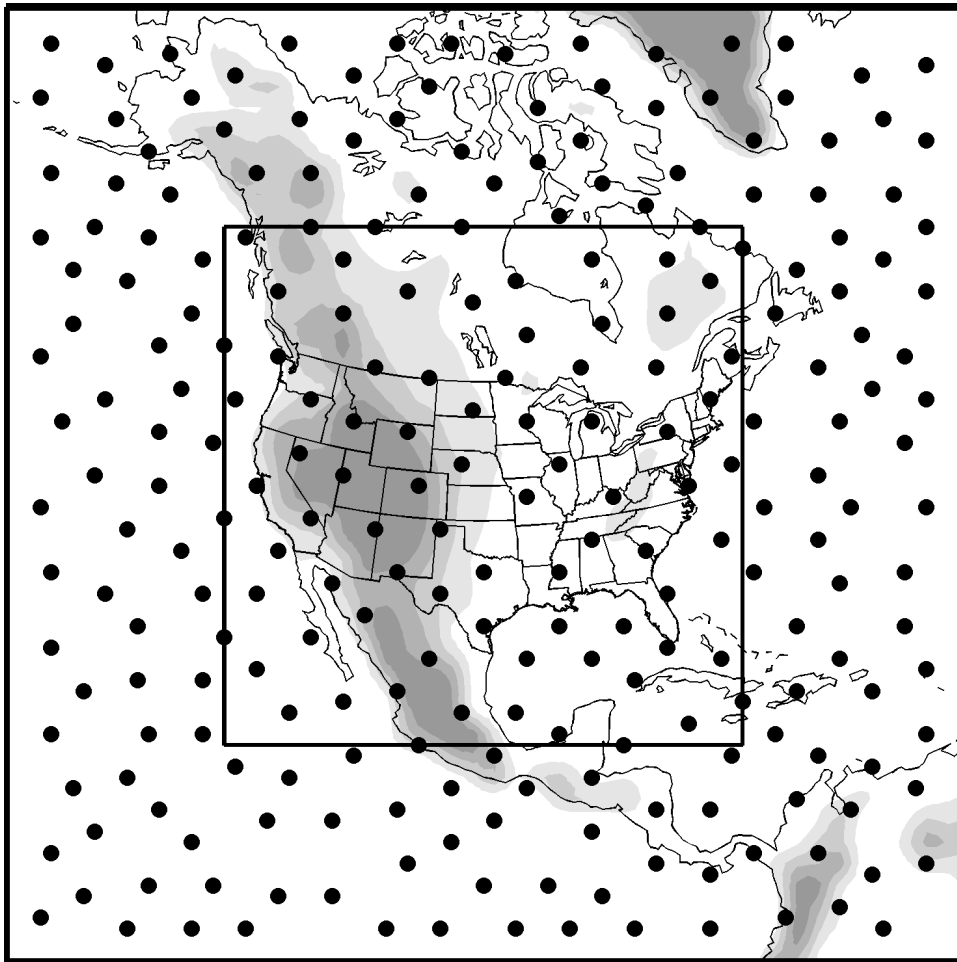


Figure 1: WRF pseudo-global domain (full grid), limited-area domain (inner box) and surface pressure observation network (dots). Shading indicates model surface elevation every 600 m.

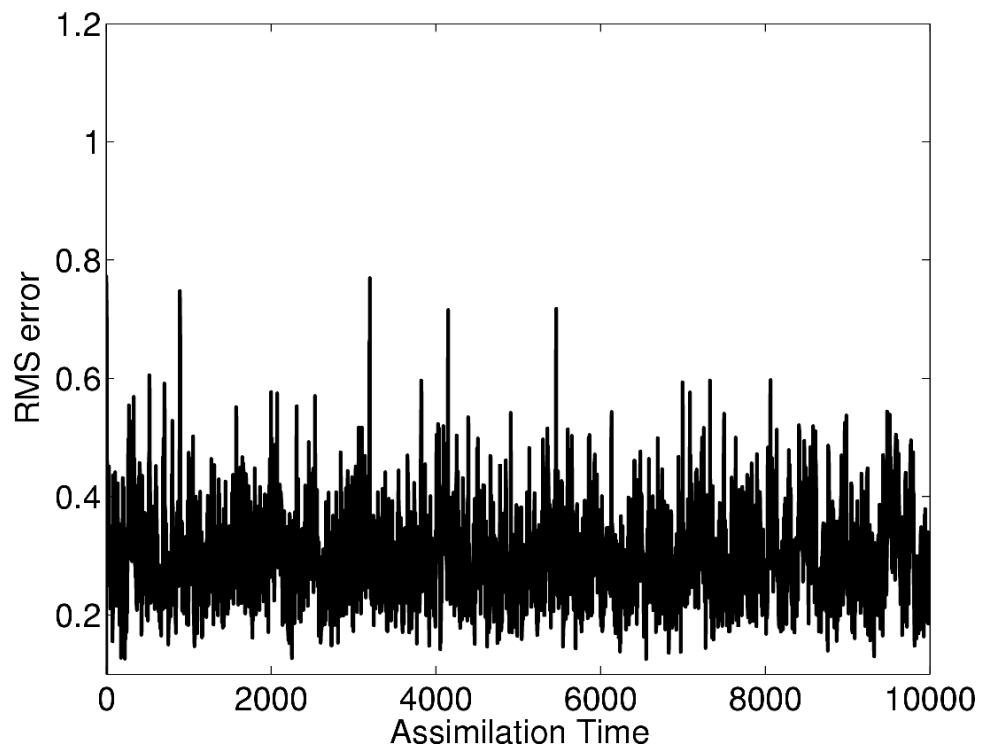


Figure 2: RMS error as a function of time (E_k) for the LE global ensemble.

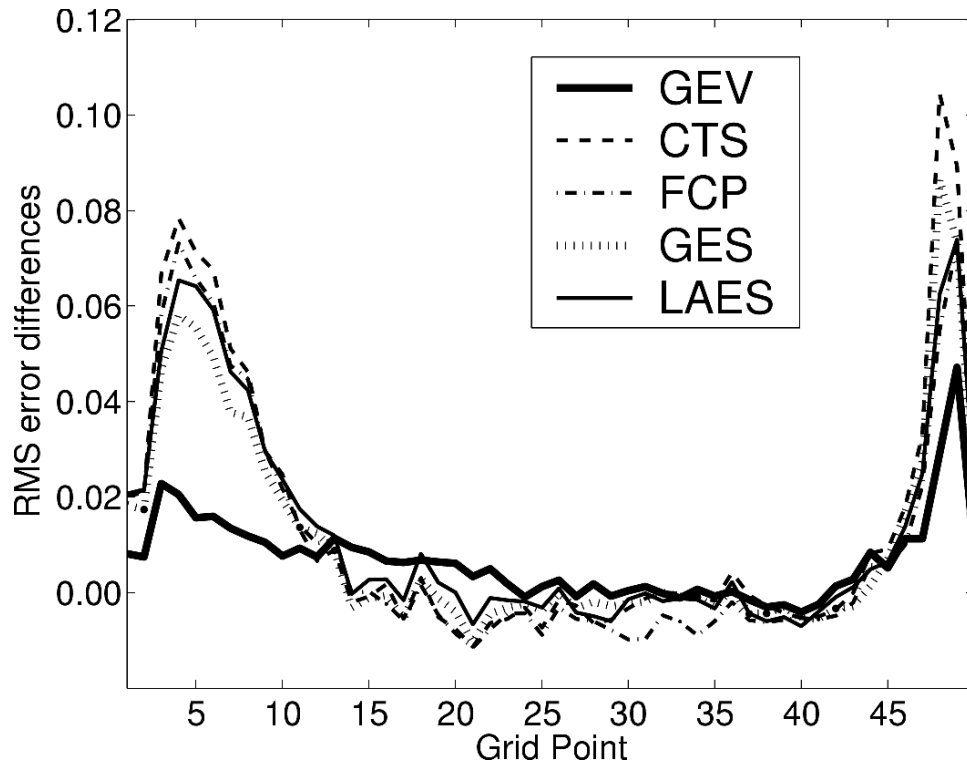


Figure 3: Difference between each method's RMS error as a function of space and global ensemble's RMS error as a function of space ($E_i^{meth} - E_i^{global}$) for all limited-area domain grid points. Positive values indicate RMS errors greater than the global ensemble.

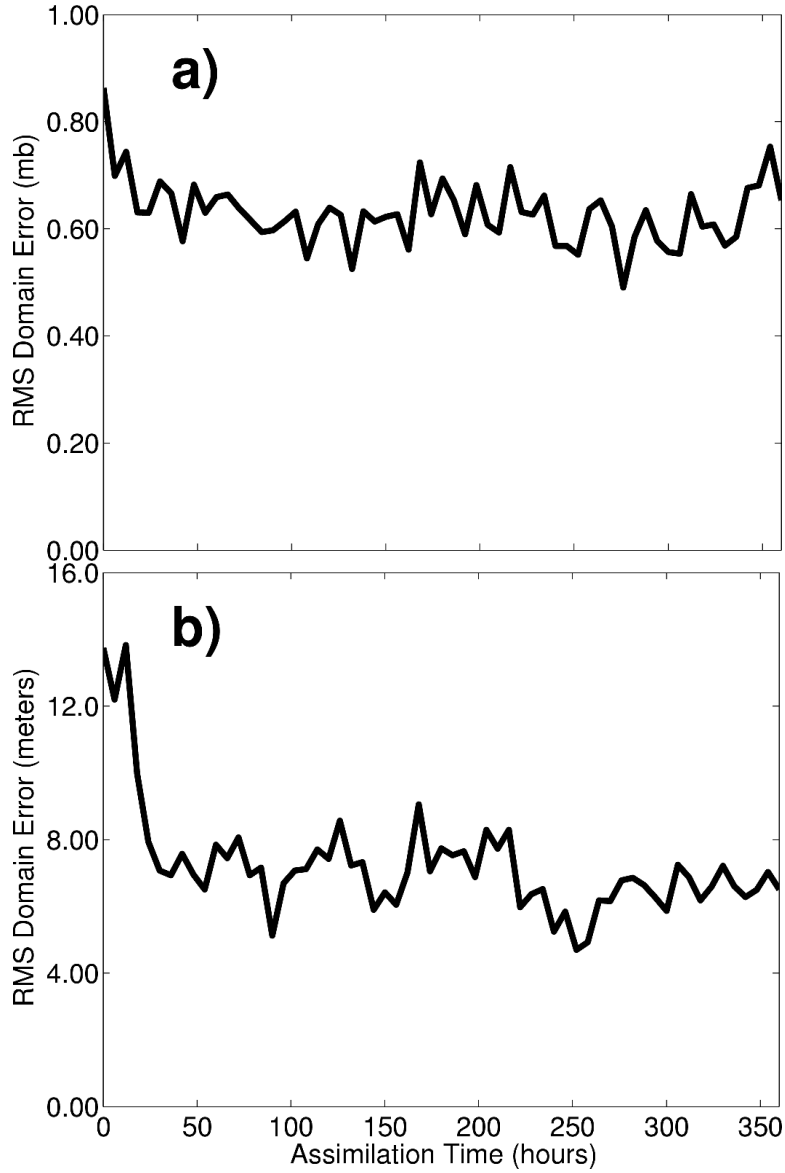


Figure 4: (a) RMS error as a function of time (E_k) for the WRF PGD dry-air mass field over all grid points in the limited-area domain. (b) As in (a) for 500 hPa geopotential height.

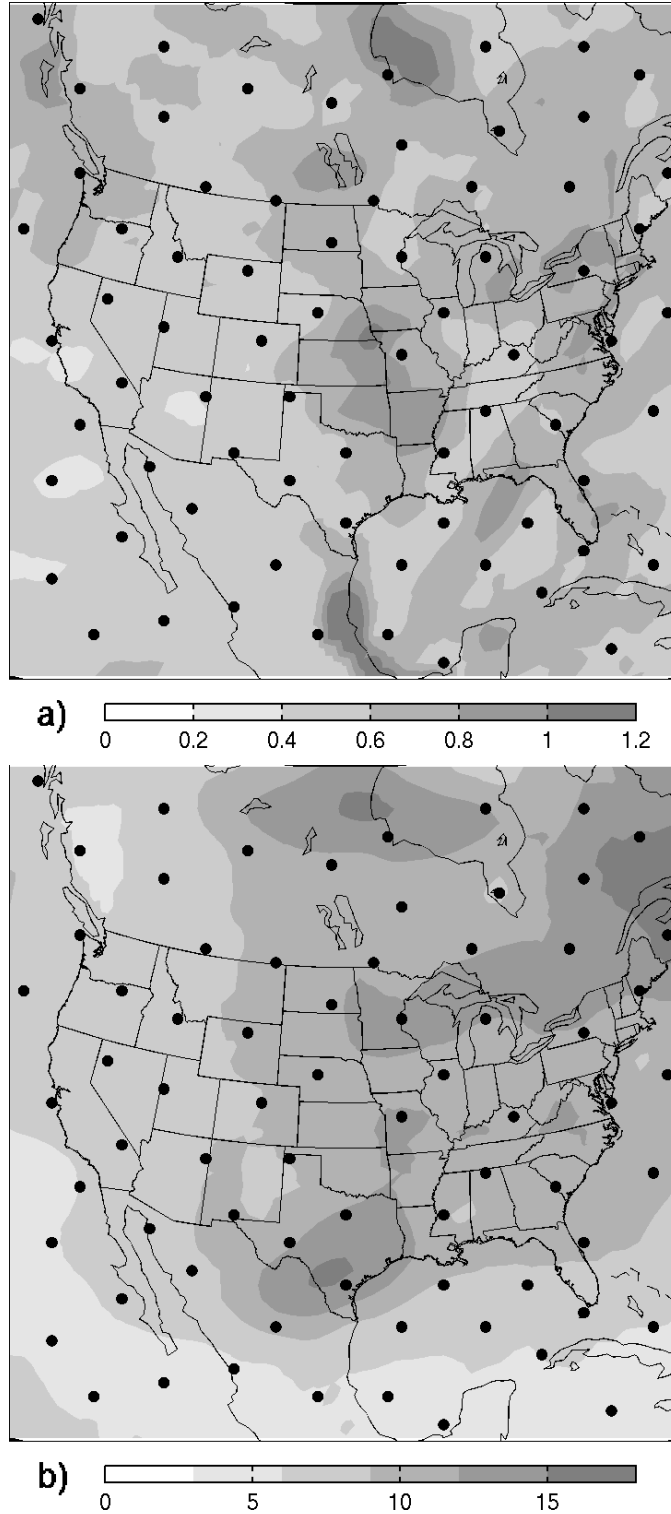


Figure 5: (a) RMS error as a function of space (E_i) in the WRF PGD dry-air mass field. Colorbar units are hPa. (b) As in (a) for 500 hPa geopotential height. Colorbar units are meters. Dots indicate observation locations.

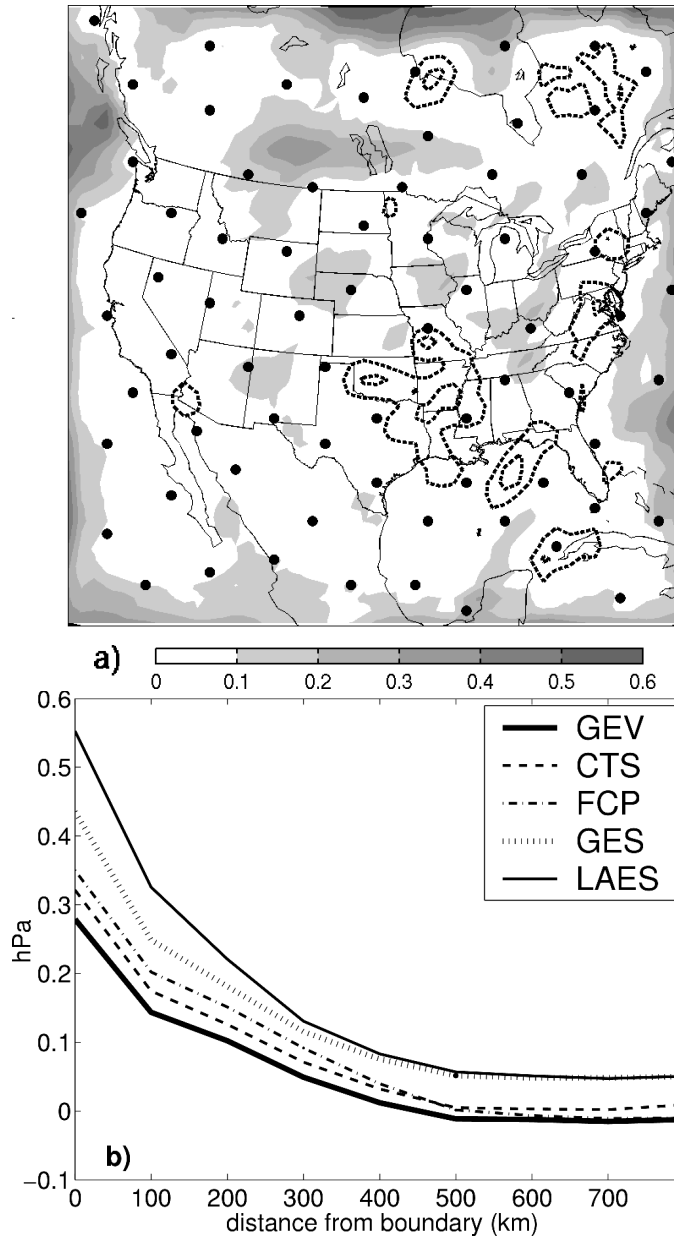


Figure 6: (a) Difference between the RMS error as a function of space in dry-air mass for the FCP method and the RMS error as a function of space for the PGD dry-air mass field at corresponding grid points in hPa. Shading indicates positive differences (i.e. larger errors than PGD), while contours denote negative differences in 0.1 hPa increments. Dots are observation locations. (b) Composite RMS error differences in each boundary method's dry-air mass field as a function of distance from the boundary.

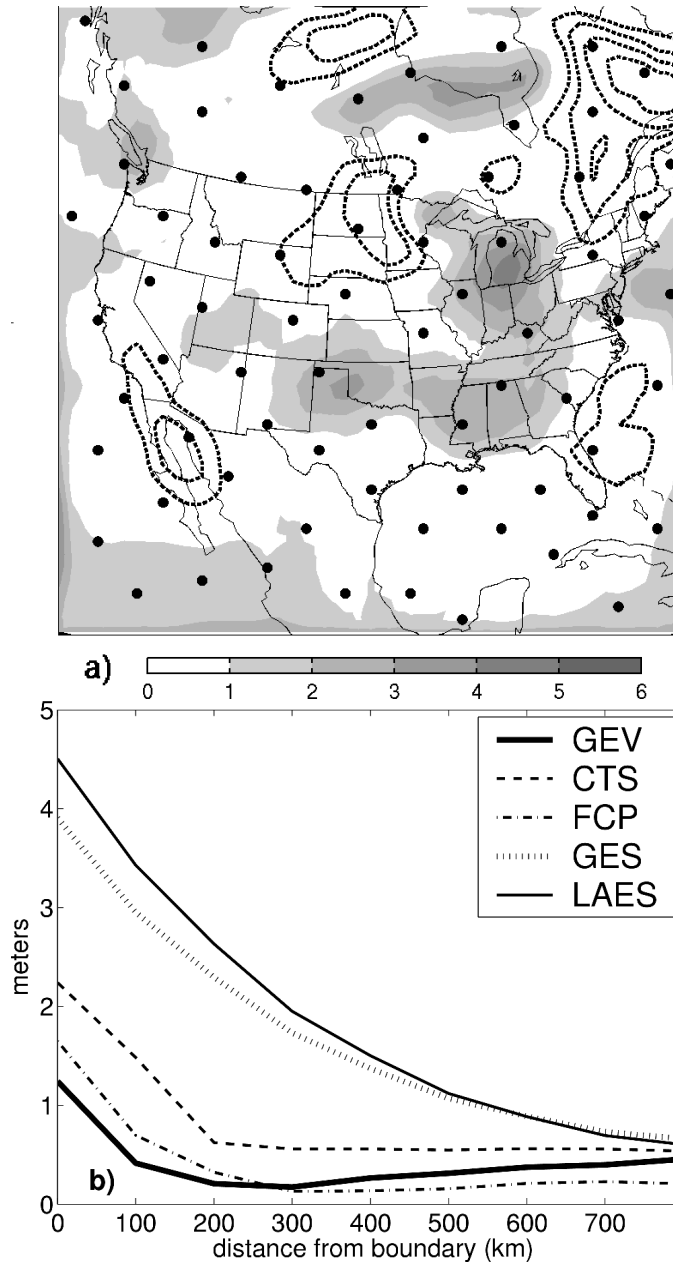


Figure 7: As in Fig. 6, but for 500 hPa geopotential height. Units are meters.